4

CRM 89-290 / March 1990

Effect of the GT Composite Requirement on Qualification Rates

Neil B. Carey

CNA A Division of Hudson Institute
CENTER FOR NAVAL ANALYSES
4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

Work conducted under contract N00014-87-C-0001.

REPORT DOCUMENTATION PAGE

Form Approved
OPM No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*

2. REPORT DATE
March 1990

3. REPORT TYPE AND DATES COVERED
Final

4. TITLE AND SUBTITLE
Effect of the GT Composite Requirement on Qualification Rates

5. FUNDING NUMBERS
C - N00014-87-C-0001
PE - 65153M
PR - C0031

6. AUTHOR(S)
Neil B. Carey

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Center for Naval Analyses
4401 Ford Avenue
Alexandria, Virginia 22302-0268

8. PERFORMING ORGANIZATION REPORT NUMBER
CRM 89-290

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
This paper analyzes whether recent changes in the definition of the Armed Forces Qualification Test (AFQT) have reduced the supply of eligible Marie Corps applicants, and whether composites other than General Technical (GT) can increase eligibility rates appreciably.

14. SUBJECT TERMS
AFQT (Armed forces qualification test), Aptitude tests, ~~ASVAB (Armed services vocational aptitude battery)~~, Enlisted personnel, Performance (human), Performance tests, Personnel selection, Qualifications, Recruits, Scoring, Statistical analysis ,

15. NUMBER OF PAGES
20

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
CPR	CPR	CPR	SAR

CNA CENTER FOR NAVAL ANALYSES
A Division of Hudson Institute 4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268 • (703) 824-2000

3 April 1990

MEMORANDUM FOR DISTRIBUTION LIST

Subj: Center for Naval Analyses Research Memorandum 89-234

Encl: (1) CNA Research Memorandum 89-234, *Concurrent Versus Enlistment ASVAB in Evaluation of New Tests*, by D.R. Divgi, Mar 1990

1. Enclosure (1) is forwarded as a matter of possible interest.

2. New kinds of tests are being evaluated as potential additions to the Armed Services Vocational Aptitude Battery (ASVAB). They are compared on the basis of the criterion variance they explain when added to the ASVAB. The evaluation may use scores on the ASVAB given during enlistment processing, or a new ASVAB may be administered concurrently with the new tests. This research memorandum compares these two research designs in terms of their effect on evaluation of new tests. The analysis uses Infantry data from the Marine Corps' Job Performance Measurement project, in which concurrent as well as enlistment ASVAB scores are available. While lower increments in explained variances are obtained when the ASVAB is concurrent, the difference between the administrations is small compared to variations across criterion variables and occupational areas.

Lewis R. Cabe
Director
Manpower and Training Program

Distribution
Reverse Page

Subj: Center for Naval Analyses Research Memorandum 89-234

Distribution List

SNDL

A1	ASSTSECNAV MRA
A1	DASN - MANPOWER(2 copies)
A2A	CNR
A6	HQMC MPR
	Attn: M
	Attn: MP
	Attn: MR
	Attn: MA (2 copies)
	Attn: MPP-39
A6	HQMC RA
A6	HQMC AVN
A6	CG MCRDAC, Washington
FF38	USNA
	Attn: Nimitz Library
FF42	NAVPGSCOL
FF44	NAVWARCOL
	Attn: E-111
FJA1	COMNAVMILPERSCOM
FJB1	COMNAVCRUITCOM
FKQ6D	NAVPERSRANDCEN
	Attn: Technical Director (Code 01)
	Attn: Director, Testing Systems (Code 63)
	Attn: Technical Library
	Attn: Director, Personnel Systems (Code 62)
	Attn: CAT/ASVAB PMO
	Attn: Manpower Systems (Code 61)
FT1	CNET
V12	CG MCRDAC, Quantico
	Attn: Director, Development Center Plans Division (Code D08) (2 copies)
	Attn: Commanding General
V12	CGMCCDC
	Attn: Training and Education Center

OPNAV

OP-01
OP-11
OP-13
OP-15

OTHER

Joint Service Selection and Classification Working Group (15 copies)
Defense Advisory Committee on Military Personnel Testing (8 copies)

CRM 89-234 / March 1990

Concurrent Versus Enlistment ASVAB in Evaluation of New Tests

D. R. Divgi

CNA A Division of Hudson Institute

CENTER FOR NAVAL ANALYSES

4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

ABSTRACT

New kinds of tests are being evaluated as potential additions to the Armed Services Vocational Aptitude Battery (ASVAB). They are compared on the basis of the criterion variance they explain when added to the ASVAB. The evaluation may use scores on the ASVAB given during enlistment processing, or a new ASVAB may be administered concurrently with the new tests. This paper compares these two research designs in terms of their effect on evaluation of new tests. The analysis uses Infantry data from the Marine Corps' Job Performance Measurement project, in which concurrent as well as enlistment ASVAB scores are available. While lower increments in explained variances are obtained when the ASVAB is concurrent, the difference between the administrations is small compared to variations across criterion variables and occupational areas.

Keywords:

EXECUTIVE SUMMARY

INTRODUCTION

The Armed Services Vocational Aptitude Battery (ASVAB) is used in selection and classification of enlisted personnel. It contains ten subtests, which measure four aptitudes—Verbal, Math, Speed, and Technical. The services have developed new tests for measuring other traits, such as psychomotor ability and spatial perception, whose measurement could also help in selection and classification. These tests are useful to the extent that they increase the predictive power of the ASVAB. The purpose of the Enhanced Computer Assisted Testing (ECAT) project is to evaluate some new tests and determine whether they should be implemented nationwide. The ECAT validation study is expected to begin by April 1990.

Estimation of the increase in predictive power requires that scores on the ASVAB, on the new test, and on a criterion be available for a group of recruits. (The criterion variable measures job performance or proficiency.) Each recruit has taken the ASVAB prior to enlistment. The criterion measure is closer in time to the new test than to the ASVAB. As a result, the predictive power of the ASVAB may be diluted, and the new test may appear more useful than it really is. Such a bias in the evaluation should be avoided as far as possible—perhaps by administering the ASVAB again, concurrently with the new test. However, this would require three more hours of testing time. Also, the recruit might not put as much effort into taking an ASVAB administered purely for research as into taking the enlistment test.

OBJECTIVE AND DATA

The objective of this paper is to determine whether enlistment and concurrent ASVAB administrations yield roughly equal values for the increase in predictive power provided by a new test. The data set was obtained from the Marine Corps' Job Performance Measurement (JPM) project. For each Marine it contains hands-on and job knowledge test scores, enlistment and concurrent ASVAB scores, and scores on four new tests. To motivate examinees to put as much effort into the concurrent ASVAB as they had put into the enlistment test, the scores became scores of record if they exceeded previous ones by a prespecified amount.

The hands-on scores are based on job tasks representative of job requirements in four military occupational specialties (MOSs) in the Infantry: Rifleman, Machinegunner, Mortarman, and Assaultman. The job knowledge tests were paper-pencil tests of information needed on the job. The new tests were Video Firing (a commercial video game), and paper-pencil tests of Space Perception, Reasoning, and Assembling Objects.

RESULTS AND CONCLUSIONS

On the whole, the new tests appeared less useful when they were added to concurrent ASVAB than to enlistment ASVAB. However, the difference was small compared to variations across MOSs, and between hands-on and job knowledge measures of job proficiency. Thus, in future research, concurrent administration of the ASVAB may be useful if some incentive is provided to motivate the examinees. Without such an incentive, the concurrent ASVAB probably will not be worth the extra testing time and expense.

The ECAT validation study of new computerized predictors includes a wide variety of jobs and of criterion variables. Its results may well show as much variability across jobs and criteria as in the present research. If they do, it will be very difficult to summarize the results, compare one test with another, and decide whether it is worthwhile to implement new tests operationally.

CONTENTS

INTRODUCTION

The Armed Services Vocational Aptitude Battery (ASVAB) is used to select and classify enlisted personnel. It contains ten subtests: General Science (GS), Arithmetic Reasoning (AR), Word Knowledge (WK), Paragraph Comprehension (PC), Numerical Operations (NO), Coding Speed (CS), Auto and Shop Information (AS), Math Knowledge (MK), Mechanical Comprehension (MC), and Electronics Information (EI). Factor analysis of the ASVAB shows that these subtests measure four factors: Verbal (GS, WK, and PC), Math (AR and MK), Speed (NO and CS), and Technical (AS, MC, and EI) [1].

There are traits the present ASVAB does not measure, such as psychomotor ability and spatial perception. The services have developed new tests for measuring such traits. Addition of new tests to the ASVAB is useful to the extent that these tests increase the predictive power of the ASVAB. Such increase is called "incremental validity." An Enhanced Computer Assisted Testing (ECAT) project is now in progress, to evaluate the incremental validities of new tests and to decide whether they should be implemented operationally. An ECAT validation study is expected to begin data collection by April 1990.

Evaluation of incremental validity requires recruits to be administered a criterion measure of job performance or proficiency. The data set must contain scores on the ASVAB, on the new test, and on the criterion for a group of recruits. First the criterion score is predicted using the ASVAB, and then the new test is added to the regression equation. The increase in the multiple correlation is the incremental validity of the new test.

Each recruit has taken the ASVAB prior to enlistment. In the ECAT validation study, the new test and the criterion will be administered after the recruit has been in a service for some time. Thus the criterion is closer in time to the new test than to the ASVAB, which may tend to reduce the predictive power of the ASVAB and hence inflate the incremental validity of the new test. This bias can be eliminated by administering the ASVAB again, along with the new test. This "concurrent" ASVAB has two shortcomings. One is that three hours of extra testing are required. The other problem is that the recruit may put less effort into it than into the enlistment ASVAB because it has no effect on his future. In such a case, the concurrent scores may have less predictive validity than those obtained under proper motivation. Therefore, it is important to know whether use of enlistment ASVAB does appreciably overestimate the incremental validity of a new test. If it does not, a concurrent ASVAB is unnecessary and its omission from the ECAT study is technically justified.

A data set from the Marine Corps' Job Performance Measurement (JPM) project can be used to compare incremental validities obtained with enlistment and concurrent ASVAB scores. The Marine Corps has developed hands-on performance tests (HOPTs) in the joint-service JPM project. An oversight committee of the National Academy of Sciences has referred to HOPTs as the "benchmark measure" of job performance ([2], p. 95). In addition to hands-on and job

knowledge tests, a concurrent ASVAB and some new tests were administered to Marines in four military occupational specialties in the Infantry: Rifleman (0311), Machinegunner (0331), Mortarman (0341), and Assaultman (0351). To motivate the examinees, the concurrent scores replaced the enlistment scores in the Marine's record if the new scores exceeded the old ones by a certain amount. The purpose of this paper is to use the JPM data set to compare enlistment and concurrent ASVAB scores in terms of the incremental validities they yield for four new predictors.

HANDS-ON PERFORMANCE MEASURES

Official Marine Corps publications and training materials, supplemented by extensive job analyses, were used to specify the domain of Infantry job requirements. Job task domains were developed for each MOS. A common core of infantry tasks required in all MOSs was identified (e.g., land navigation, tactical measures, first aid, grenade launcher), in addition to MOS-specific requirements (e.g., Rifleman, M16A2 rifle; Machinegunner, M60 machinegun; Mortarman, 60-mm and 81-mm mortars; and Assaultman, Dragon and SMAW). Tasks were sampled from each domain so that hands-on test scores would generalize to the full range of Infantry job requirements within that domain. Scores were computed for both the common core and the MOS-specific components and then weighted to create a hands-on total score for each MOS. Details are provided in [3].

Because of the uniqueness of the MOS-specific components, HOPT scores for the different MOSs were not on the same scale. To achieve comparability, the HOPTs were linearly transformed to a standard scale as follows: Each HOPT was regressed on time in service (TIS), its square (TIS_SQ) and all ten enlistment ASVAB subtests, separately for each MOS. Using the regression in a given MOS, the expected HOPT score at TIS of 24 months was computed for each person in the 1980 Reference Population [4]. A random normal error term with a standard deviation equal to the standard error of estimate was added to generate a simulated HOPT score for each person. The mean and standard deviation of the simulated HOPT scores were calculated over the Reference Population. Then, using this mean and standard deviation, the HOPT scale was transformed so that its mean and standard deviation in the Reference Population were 50 and 10 (except for the sampling error caused by the random error term). Because the population variance is 100, incremental variances explained by new tests are directly interpretable as percentages of the total variance.

OTHER TESTS

Paper-pencil job knowledge tests (JKTs), which asked for information about tasks in the HOPT, were also administered. JKT scores were standardized in the same way as HOPT scores.

Four new predictors were evaluated: Video Firing, Space Perception, Reasoning, and Assembling Objects. Video Firing assessed psychomotor ability with a computer game of firing at a target on a television screen. The other three tests were paper-pencil. Space Perception tested spatial visualization with items that involved drawings of folded and unfolded pieces of

paper. The Reasoning test measured spatial reasoning and pattern recognition. The Assembling Objects test measured spatial visualization and mental rotation.

SUBJECTS

The subjects were active-duty male Marines with various durations of service in four Infantry MOSs: Rifleman (0311), Machinegunner (0331), Mortarman (0341), and Assaultman (0351). The data were collected as part of the Marine Corps' Job Performance Measurement project. Subjects were administered the ASVAB during the study. This will be referred to as the concurrent ASVAB. In addition, enlistment ASVAB scores were available from the subjects' files. Only those Marines with complete data (i.e., criterion, new predictor, and ASVAB scores), and no more than 48 months of service, were included in the analyses. The useful sample consisted of 864 Riflemen, 234 Machinegunners, 223 Mortarmen, and 251 Assaultmen.

METHODOLOGY

Preliminary analyses showed that results were practically the same whether the subtest scores or factor scores on the ASVAB were used. Therefore, in the interests of parsimony, scores on the four factors were used. Hands-on and job knowledge criteria were analyzed separately.

Time in service has a major influence on performance scores because it represents training while on the job. Analyses show that, as the recruit learns more, the learning rate decreases. Therefore, in each regression, TIS and TIS_SQ were first entered into the prediction equation. Once this was done, the regression results for predictors added subsequently indicated what would have happened if all subjects in the study had the same length of service.

There is another way in which TIS may affect the prediction of criterion scores. As time passes, a Marine's true ASVAB scores may change somewhat. As a result, the regression coefficients for the enlistment factor scores may decrease as TIS increases. For each of the four factors, this effect was allowed for by adding a new predictor equal to the product of the factor score with TIS. For the Verbal factor this new variable was called T_VERBAL, and so on.

A separate analysis was performed for each new test. In each MOS, the four ASVAB factors and then the new test were entered into the regression. The resulting residual variance was compared with that using only the four factors. (Residual variance equals the square of the standard error of estimate.) The reduction in residual variance due to addition of the new test was the variance explained by the test. Such calculations automatically incorporate the standard correction used in adjusting the multiple correlation for its upward bias [5]. (Because of this correction, when the reduction in residual variance is smaller than that expected from pure chance, the estimated value comes out negative.)

The data were collected in a sample of recruits who have been selected using the ASVAB. Therefore, validity of the ASVAB is lower in this selected sample than in the national population. Corrections for such restriction of range are simple for the quantities of interest in this paper. First only the ASVAB, with or without slope changing with TIS, is used in the regression equation. When the residual variance of this analysis is subtracted from 100, one obtains the variance explained by the ASVAB in the national population. Then one new test is added as a predictor, and the additional variance explained by this test is calculated. According to the assumptions used in range correction, this variance is the same in the unselected national population as in the selected recruit population.

RESULTS

Since recruits are a selected sample, their ASVAB scores have a smaller spread than in the national population. The degree of this range restriction can be quantified in different ways. For the purposes of this paper, the relevant ASVAB "score" is the predicted value of the criterion variable, using the enlistment ASVAB and holding TIS constant. The spread of this score in the recruit sample, relative to the national population, is described by the ratio of standard deviations in these two groups. For the hands-on criterion, this ratio was .634 in MOS 0311, .654 in MOS 0331, .785 in MOS 0341, and .696 in MOS 0351. For job knowledge the corresponding ratios were .526, .562, .686, and .542. Thus, MOS 0341 is less restricted than the others.

Table 1 shows percentages of population variance explained by the three ways of using ASVAB scores. E_ASVAB in the Predictor column means that the enlistment ASVAB factors were used in the regression. C_ASVAB means that concurrent ASVAB factors were used. E_ASVAB* means that addition of T_VERBAL, etc., allowed regression weights to change with TIS. Addition of this interaction term turned out to have a noticeable effect only on the hands-on criterion in MOS 0341.

Table 1. Percent variances in population explained by enlistment ASVAB (E_ASVAB), by enlistment ASVAB plus interaction with TIS (E_ASVAB*), and by concurrent ASVAB (C_ASVAB)

		MOS			
Criterion	Predictor	0311	0331	0341	0351
HOPT	E_ASVAB	44.7	54.8	41.2	33.9
HOPT	E_ASVAB*	44.8	54.5	44.5	34.1
HOPT	C_ASVAB	45.6	55.7	40.6	39.8
JKT	E_ASVAB	65.2	59.2	50.9	55.6
JKT	E_ASVAB*	65.1	59.1	52.2	56.5
JKT	C_ASVAB	67.8	65.3	56.9	61.3

Table 2 shows incremental percentages of explained variance due to the four new predictors. Entries in the Predictor column have the same meanings as in table 1. Again, in most cases, the numbers are affected very little by allowing ASVAB regression weights to change with TIS.

Table 2. Percent variances explained by new tests when added to enlistment ASVAB (E_ASVAB), to enlistment ASVAB plus interaction with TIS (E_ASVAB*), and to concurrent ASVAB (C_ASVAB)

		New Test			
Criterion	Predictor	Firing	Space	Reasoning	Objects
MOS 0311: Rifleman, N = 864					
HOPT	E_ASVAB	2.27	0.52	0.87	1.26
HOPT	E_ASVAB*	2.33	0.55	0.89	1.27
HOPT	C_ASVAB	2.09	0.35	0.46	0.94
JKT	E_ASVAB	0.13	0.22	1.49	2.44
JKT	E_ASVAB*	0.13	0.22	1.49	2.41
JKT	C_ASVAB	0.07	0.02	0.62	1.48
MOS 0331: Machinegunner, N = 234					
HOPT	E_ASVAB	0.90	2.56	0.80	0.70
HOPT	E_ASVAB*	0.87	2.64	0.72	0.63
HOPT	C_ASVAB	1.14	3.00	1.29	0.65
JKT	E_ASVAB	0.11	1.09	0.26	3.95
JKT	E_ASVAB*	0.09	1.20	0.27	4.11
JKT	C_ASVAB	0.11	0.99	0.16	2.76
MOS 0341: Mortarman, N = 223					
HOPT	E_ASVAB	0.56	0.91	0.76	3.25
HOPT	E_ASVAB*	0.49	0.75	0.79	2.59
HOPT	C_ASVAB	0.26	0.68	−0.08	2.46
JKT	E_ASVAB	1.36	1.69	4.69	3.78
JKT	E_ASVAB*	1.10	1.49	4.66	3.41
JKT	C_ASVAB	0.66	1.24	2.09	2.62
MOS 0351: Assaultman, N = 251					
HOPT	E_ASVAB	1.72	0.88	0.14	1.19
HOPT	E_ASVAB*	1.87	0.97	−0.01	1.12
HOPT	C_ASVAB	0.67	0.27	−0.14	0.50
JKT	E_ASVAB	0.03	0.50	0.45	1.97
JKT	E_ASVAB*	0.08	0.52	0.32	1.84
JKT	C_ASVAB	-0.15	0.24	−0.05	0.70

The primary comparison is between the E_ASVAB* and C_ASVAB rows for HOPT within each MOS. Of the 16 comparisons, C_ASVAB yields smaller increments in 12 cases. However, the difference is usually a fraction of a percent, and hence much less important than the difference between one MOS and another. Differences are larger for JKT.

The additional variance explained by a given subtest depends on the criterion. Video Firing and Space Perception contribute more to HOPT than to JKT; the opposite is true of Reasoning and Assembling Objects. Variations are found across MOSs as well: Video Firing makes its largest contribution in the Rifleman MOS, Space Perception in the Machinegunner MOS, and Reasoning and Object Assembly in the Mortarman MOS. Such variations make it very difficult to compare the tests.

CONCLUSIONS

In conclusion, as a matter of research design, a concurrent ASVAB does not appear to be important for avoiding an upward bias in evaluation of new tests, particularly if hands-on performance is the primary criterion variable. The difference between incremental variances using concurrent and enlistment ASVABs is minor compared to variations across different criterion variables and occupational specialties. In addition, it should be remembered that Marines in the JPM study had an incentive to do well on the concurrent ASVAB—a chance to increase their scores of record. Without such an incentive to improve motivation, a concurrent ASVAB may yield misleading results.

The changes in incremental variance across different MOS and from HOPT to JKT should be seen as a danger signal. The research design for the ECAT validity study of new computerized tests is given by Wolfe [6]. This study covers 31 different occupations. The criterion variables include hands-on performance, supervisor ratings, job knowledge tests, performance on simulators and in laboratories (when available), and training school grades. Different criteria will be available in different occupations. If results of this study show the kinds of variations found in table 2, it will be very difficult to summarize them and to conclude that one test is more useful than another.

REFERENCES

[1] CNA Memorandum 83-3135, *A Factor Analysis of ASVAB Form 8a in the 1980 DOD Reference Population*, by Peter H. Stoloff, Aug 1983 (05833135)[1]

[2] Alexandra K. Wigdor and Bert F. Green, Jr., eds. *Assessing the Performance of Enlisted Personnel: Evaluation of a Joint Service Research Project.* Washington, D.C.: National Academy Press, 1986

[3] CNA Research Contribution 570, *Developing a Competency Scale for Hands-on Measures of Job Proficiency*, by Paul W. Mayberry, Dec 1987 (02057000)

[4] CNA Report 116, *The ASVAB Score Scales: 1980 and World War II*, by Milton H. Maier and William H. Sims, Jul 1986 (94011600)

[5] R. J. Wherry, "A New Formula for Predicting the Shrinkage of the Coefficient of Multiple Correlation." *Annals of Mathematical Statistics* (1931): 440-457

[6] John H. Wolfe, "Design for Joint Service Validation of New Predictors," briefing presented to the Defense Advisory Committee on Military Personnel Testing, 28 Jul 1989

1. The number in parentheses is a CNA internal control number.